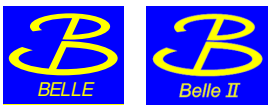
Published for SISSA by Springer

Received: *July 3, 2025*
Accepted: *November 7, 2025*
Published: *December 15, 2025*

# Search for an axion-like particle in $B \to K^{(*)} a(\to \gamma\gamma)$ decays at Belle




## The Belle and Belle II collaborations

*Full author list at the end of the paper*

*E-mail:* coll-publications@belle2.org, yjkwon63@yonsei.ac.kr

Abstract: We report a search for an axion-like particle $a$ in $B \to K^{(*)} a$ decays using data collected with the Belle detector at the KEKB asymmetric-energy electron-positron collider. The search is based on a 711 fb$^{-1}$ data sample collected at the $\Upsilon(4S)$ resonance energy, corresponding to a sample of $772 \times 10^6$ $\Upsilon(4S)$ events. In this study, we search for the decay of the axion-like particle into a pair of photons, $a \to \gamma\gamma$. We scan the two-photon invariant mass in the range 0.16 GeV–4.50 GeV for the $K$ modes and 0.16 GeV–4.20 GeV for the $K^*$ modes. No significant signal is observed in any of the modes, and 90% confidence level upper limits are established on the coupling to the $W$ boson, $g_{aW}$, as a function of $a$ mass. The limits range from $3 \times 10^{-6}$ GeV$^{-1}$ to $3 \times 10^{-5}$ GeV$^{-1}$, improving the current constraints on $g_{aW}$ by a factor of two over the most stringent previous experimental results.

Keywords: Beyond Standard Model, B Physics, $e^+$-$e^-$ Experiments

ArXiv ePrint: 2507.01249



https://doi.org/10.1007/JHEP12(2025)109

Modified publication 2026

## Contents



## 1 Introduction

The Peccei-Quinn (PQ) theory introduces axions as a solution to the strong-CP problem, positioning them as promising candidates for dark matter in extensions to the standard model (SM) [1–3]. Other extensions to the SM introduce axion-like particles (ALPs) $a$, which share the quantum numbers of axions, but, unlike the QCD axion, have couplings that are independent of their masses. While the PQ axion is expected to have a mass below $\mathcal{O}(1\,\text{MeV})$ inversely proportional to its decay constant $f_a$ ($m_a \approx 6.3\,\text{eV} \cdot 10^6\ \text{GeV}/f_a$), ALPs encompass a broader mass spectrum. ALPs have the potential to address fundamental problems [4–6]. In particular, they can act as dark matter mediators through the axion portal at the $\mathcal{O}(\,\text{GeV})$ scale [7].

Recent years have seen a surge of interest in ALPs in the MeV and GeV mass range [8–19]. The ALP couplings to photons, leptons, and gluons have been extensively studied by collider and beam dump experiments [20–27]. In contrast, the coupling to $W^\pm$ bosons remains relatively uncharted [28]. The most recent study of ALP coupling to $W^\pm$ bosons was conducted by the *BABAR* experiment using the $B^+ \to K^+ a(\to \gamma\gamma)$ decay [29]. In this paper, we work in natural units ($\hbar = c = 1$), and charge-conjugate modes are implicitly included. Although the ALP is assumed to primarily couple to the $W$ boson, gauge-boson mixing induces a coupling to photons, resulting in a nearly 100% branching fraction for the $a \to \gamma\gamma$ decay in the analysed mass range, $m_a < M_W$.

The coupling of the ALP to $W^\pm$ bosons is described by the Lagrangian [28],

$$\mathcal{L} = -\frac{g_{aW}}{4} a W_{\mu\nu} \tilde{W}^{\mu\nu}, \tag{1.1}$$



JHEP12(2025)109

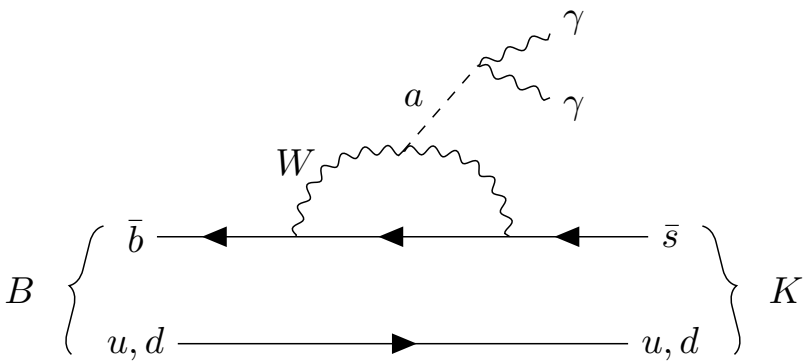



**Figure 1.** The Feynman diagram for a $B \to K^{(*)}a(\to \gamma\gamma)$ decay.

where $a$ is the ALP field, $g_{aW}$ is the coupling strength of the $a$ to $W$ bosons, and $W_{\mu\nu}$ is the gauge boson field strength for a $W$ boson, with dual tensor $\tilde{W}^{\mu\nu} = \epsilon^{\mu\nu\rho\sigma}W_{\rho\sigma}/2$, where $\epsilon^{\mu\nu\rho\sigma}$ is the four-dimensional Levi-Civita symbol. The branching fraction of the process $B \to K^{(*)}a(\to \gamma\gamma)$, which is shown in figure 1, depends quadratically on $g_{aW}$ (see the details in the appendix A), reflecting the fact that the ALP decay width is given by

$$\Gamma = \frac{1}{\tau_a} = g_{aW}^2 m_a^3 \sin^4\theta_W/64\pi\,, \tag{1.2}$$

where $\theta_W$ is the weak mixing angle, and $\tau_a$ is the ALP lifetime [29].

In this paper, we report a search for ALPs in $B \to K^{(*)}a(\to \gamma\gamma)$ decays using four kaon modes, $K_S^0$, $K^+$, $K^{*0}$, and $K^{*+}$. The data were collected with the Belle detector [30] at the asymmetric-energy $e^+e^-$ KEKB collider [31], operating at a centre-of-mass energy of 10.58 GeV, corresponding to the $\Upsilon(4S)$ resonance in $e^+e^-$ collisions. This search is based on a data set of $772 \pm 11$ million $\Upsilon(4S)$ mesons, corresponding to an integrated luminosity of 711 fb$^{-1}$. Our study has better sensitivity than the previous *BABAR* study, leveraging the higher total integrated luminosity of the Belle experiment, as well as additional kaon modes. The ALP mass hypotheses for the $K^+$ and $K_S^0$ modes range from 0.16 GeV to 4.50 GeV, while for the $K^{*+}$ and $K^{*0}$ modes the range is 0.16 GeV to 4.20 GeV. This analysis also probes regions of parameter space in which the ALP lifetime is non-negligible, leading to displaced decays a few centimetres from the interaction point. Once we obtain the results separately for each mode, we combine the four kaon modes using a simultaneous fit to improve the constraint on $g_{aW}$.

## 2 The Belle detector and simulation

The Belle detector is a general purpose detector, described in detail in [30]. It has a cylindrical symmetry around the beam line, with the $z$-axis being defined as the direction opposite to the positron beam. The detector consists of six subdetectors: a silicon vertex detector (SVD) for precise vertex determination, a central drift chamber (CDC) for reconstruction of charged particle trajectories (tracks) and for measuring their momentum, an aerogel Cherenkov counter (ACC) and a time-of-flight scintillation counter (TOF) for particle identification, an



JHEP12(2025)109

electromagnetic calorimeter (ECL) for photon detection, photon energy measurement, and electron identification, surrounded by a 1.5 T superconducting solenoid, and resistive plate chambers installed in the flux-return yoke to detect $K_L^0$ and $\mu$ (KLM). The ECL, which is crucial for this analysis, consists of 8736 CsI(Tl) crystals with a nearly projective geometry covering the polar angle range of $12° < \theta < 157°$.

To mitigate possible biases in the analysis, we establish the event selection and the search method using simulation, and we validate these with control modes and off-resonance data before the experimental data are examined. The signal $B \to K^{(*)}a(\to \gamma\gamma)$ processes are simulated using the EvtGen generator [32]. We generate 60 samples for $K$ modes, with $a$ masses ranging from 0.16 GeV to 4.50 GeV, and 56 samples for $K^*$ modes, with $a$ masses ranging from 0.16 GeV to 4.20 GeV. The long-lived $a$ sample is simulated with the ALP lifetime $c\tau$ varying from 10 mm to 500 mm, where $\tau$ is the proper lifetime. In addition, we simulate a sample with uniformly distributed diphoton invariant mass in the entire mass range from 0.01 GeV to 4.78 GeV for $K$ modes, and 0.01 GeV to 4.28 GeV for $K^*$ modes with 10 MeV intervals.

Several processes contribute as background, including $e^+e^- \to \Upsilon(4S) \to B\overline{B}$ with $B$-meson decaying to SM particles that can mimic the signal, as well as $e^+e^- \to q\overline{q}$ ($q = u, d, s, c$) continuum processes. We simulate the background processes $\Upsilon(4S) \to B\overline{B}$ and $e^+e^- \to q\overline{q}$ ($q = u, d, s, c$) using the EvtGen [32] and PYTHIA [33] generators. Final-state radiation is simulated using PHOTOS [34]. The detector response is simulated with GEANT3 [35]. Both experimental data and simulated events are converted to the Belle II format using B2BII [36] and then analysed using the Belle II analysis software framework [37, 38].

## 3 Selection of signal events

The charged particles in $B \to K^{(*)}a(\to \gamma\gamma)$ decays, which originate from one of the four kaon modes $K^+$, $K^{*0}$, $K^{*+}$, and $K_S^0$, are required to have a point of closest approach to the interaction point of less than 4 cm in the $z$-direction and less than 3 cm in the radial direction. The charged kaons ($K^+$) are identified using a likelihood ratio $\mathcal{P}$, which compares two particle hypothesis $i$ and $j$, $\mathcal{P}(i : j) = \mathcal{L}_i/(\mathcal{L}_i + \mathcal{L}_j)$. The likelihoods $\mathcal{L}$ are calculated based on the Cherenkov photon yield in the ACC, the energy-loss measurements in the CDC, and the time-of-flight information from the TOF. To select $K^+$, we require $\mathcal{P}(K : \pi) > 0.6$ and $\mathcal{P}(K : p) > 0.4$. The $K^*$ candidates are reconstructed from the decay modes $K^{*0} \to K^+\pi^-$ and $K^{*+} \to K_S^0\pi^+$ within the invariant mass range of 0.8 GeV to 1.0 GeV. To select charged pions $\pi^\pm$ we require $\mathcal{P}(\pi : K) > 0.4$ and $\mathcal{P}(\pi : p) > 0.7$, with the likelihood ratio definitions being the same as for the charged kaon. The identification efficiencies for pions and kaons are 96.9% and 85.3%, respectively, with misidentification rates for pions as kaons and vice versa of 10.3% and 1.3%, respectively, while the misidentification rates for protons are negligible. The $K_S^0$ candidates are reconstructed from the $K_S^0 \to \pi^+\pi^-$ decay mode. Pairs of oppositely charged pions are combined, and an artificial neural network [39] is then used to identify the $K_S^0$ candidates. The neural network is trained with 13 variables. The dipion invariant mass is required to be within a range of 20 MeV around the nominal mass of the $K_S^0$ [40]. This range corresponds to approximately $\pm 5\sigma$ window around $K_S^0$ mass, where $\sigma$ is the $\pi^+\pi^-$ invariant mass resolution. Detailed information is provided in ref. [41]. Particles with high



JHEP12(2025)109

lepton probability are excluded from our analysis. For the electron identification, we primarily use the information from the ECL along with other subdetectors. The electron likelihood ratio is defined as $\mathcal{P}(e) = \mathcal{L}_e/(\mathcal{L}_e + \mathcal{L}_{\bar{e}})$, where $\mathcal{L}_e$ is the electron likelihood and $\mathcal{L}_{\bar{e}}$ is the non-electron likelihood, a product of likelihoods from the ACC, CDC, and ECL. For muon identification, the likelihood is calculated using the information from the KLM. The likelihood ratio is defined as $\mathcal{P}(\mu) = \mathcal{L}_\mu/(\mathcal{L}_\mu + \mathcal{L}_\pi + \mathcal{L}_K)$, where the value $\mathcal{L}_\mu$ is determined based on whether the charged particle has an associated KLM signature. We require $\mathcal{P}(e) < 0.9$ and $\mathcal{P}(\mu) < 0.9$ to veto leptons.

The ALP candidates are reconstructed from the $a \to \gamma\gamma$ decay. Photons are identified from ECL energy deposits that are not associated with reconstructed charged particles. To suppress the contribution from photons that originate from beam background, each photon candidate is required to have a minimum energy that depends on the ECL region: $E_\gamma > 50\,\text{MeV}$ in the barrel ($32.2° < \theta < 128.7°$), $E_\gamma > 100\,\text{MeV}$ for the forward endcap ($12.0° < \theta < 31.4°$), and $E_\gamma > 150\,\text{MeV}$ for the backward endcap ($131.5° < \theta < 157.1°$), where $\theta$ is the polar angle of the photon candidate in the laboratory frame.

The signal $B$ candidates are then formed by combining an ALP candidate with a $K^{(*)}$ candidate. We select $B$ candidates using two kinematic variables: the beam-constrained mass $M_{\text{bc}} = \sqrt{E^2_{\text{beam}} - p^2_B}$ and the energy difference $\Delta E = E_B - E_{\text{beam}}$, where $p_B$ and $E_B$ are the momentum and energy of the $B$ candidate in the centre-of-mass (c.m.) frame, and $E_{\text{beam}}$ is the beam energy in the c.m. frame. We require $M_{\text{bc}} > 5.27\,\text{GeV}$ and $-0.2\,\text{GeV} < \Delta E < 0.1\,\text{GeV}$. A kinematic fit is applied to the selected $B$ candidates, constraining the $B$ decay position to the interaction point, and the invariant mass of the $K^{(*)}\gamma\gamma$ to the nominal $B$ mass. After $B$ meson reconstruction, about 25% of events have multiple signal candidates. For such events, we select the candidate with the smallest $|\Delta E|$. According to Monte Carlo simulations, approximately 85% of the candidates are correctly reconstructed after the candidate selection.

A series of fast boosted decision tree (BDT) classifiers [42] are used after event selection to suppress the background processes. To train the BDT classifiers, ALP simulations with uniformly distributed diphoton invariant mass, $M_{\gamma\gamma}$, are used. This ensures that the classifiers do not learn or exploit the signal peak position in their background rejection strategy, and consequently avoids bias in the search for ALPs at a specific mass hypothesis. Since the distributions of input variables differ significantly between the low and high mass regions, the data above and below 1 GeV are treated separately to enhance classification power. The training of all classifiers is based on simulated samples that are statistically independent of those used to develop and validate the fitting strategy. This separation ensures an unbiased optimization of the background suppression procedure before its application to the unblinded data.

We apply two continuum suppression ($CS$) classifiers to suppress the dominant background process, $e^+e^- \to q\bar{q}$. The first classifier, $CS1$, uses 10 variables. These are the ratio of zeroth to second Fox-Wolfram moments ($R_2$) [43], several modified Fox-Wolfram moments [44, 45], the cosine of the angle between the signal $B$-thrust and the rest-of-event (ROE) thrust axes, the cosine of the angle between the signal $B$-thrust and the $z$ axes, and the magnitude of the ROE thrust. The thrust axis is defined as the axis that maximizes the sum of the projected momenta of all particles [46]. The signal $B$ thrust axis is calculated using



JHEP12(2025)109

particles in the signal candidate, while the ROE thrust axis is determined from the charged particles and photons not used in the signal candidate reconstruction. The second classifier, $CS2$, is trained using events with $CS1 > 0.1$ and utilizes the event sphericity and aplanarity (which is a linear combination of the sphericity eigenvalue and 3/2 of the third sphericity eigenvalue [47]), the sum of the absolute value of the momenta of the particles moving along or against the thrust axis, harmonic moments (coefficients of the spherical harmonic event expansion around the thrust axis), the energy asymmetry between the two photons, modified Fox-Wolfram moments, and number of ALP candidates per event ($N_{\text{cand}}$). We find, based on our simulation study, that $N_{\text{cand}}$ is larger for signal events than for background ones. The most discriminating variables in both classifiers are the cosine of the angle between the signal $B$-thrust and ROE-thrust axes for the low ALP mass hypothesis and the reduced Fox-Wolfram $R2$ for the high ALP mass hypothesis.

A major background arises from $\pi^0$ mesons produced in $\Upsilon(4S) \to B\bar{B}$ or $e^+e^- \to q\bar{q}$ processes, decaying into $\gamma\gamma$. To suppress these, we first calculate identification variables by combining one photon, $\gamma_a$, from the reconstructed ALP candidate with any other photon, $\gamma_b$, in the ROE. For each pair of photons $(\gamma_a, \gamma_b)$, we perform three binary classification tests, each based on different assumptions about their source:

(1) We compare a pair in which $\gamma_a$ is from a true ALP signal and $\gamma_b$ is from the ROE in a signal event, against a pair from the same $\pi^0$ decay in a background event;

(2) We compare a pair in which $\gamma_a$ is from a true ALP signal and $\gamma_b$ is from the ROE in a signal event, against a pair where $\gamma_a$ and $\gamma_b$ are from different processes in a background event;

(3) We compare a pair in which both $\gamma_a$ and $\gamma_b$ come from different processes in a background event, against a pair where both photons come from the same particle decay, such as a $\pi^0$, also in a background event.

The tests are carried out with BDTs trained with multiple variables from Monte Carlo simulated events such as the ratio of energies $E_9$ and $E_{25}$ in the inner $3 \times 3$ and $5 \times 5$ crystals around the central crystal ($E_9/E_{25}$), the energy of the most energetic crystal in the ECL cluster, the sum of weights of all crystals in an ECL cluster, the photon energy, as well as kinetic properties of the diphoton system such as its mass, energy, opening angle, transverse momentum, and energy asymmetry. Among the various possible partner photons $\gamma_b$, the one that yields the highest $\pi^0$-like score in each test is selected. These three individual scores are then combined into a single score, which is used to calculate the probabilities, $P_{\pi^0}(\gamma_1)$ and $P_{\pi^0}(\gamma_2)$, for each signal candidate photon, $\gamma_1$ and $\gamma_2$, to originate from a $\pi^0$.

For each ALP $a$ mass range and kaon mode, selection criteria on the four BDT classifier scores, $CS1$, $CS2$, $P_{\pi^0}(\gamma_1)$ and $P_{\pi^0}(\gamma_2)$, are optimized using the Punzi figure of merit (PFM) [48]. A four-dimensional grid search with a step size of $\mathcal{O}(10^{-3})$ is performed, selecting the points with the highest PFM value as the minimum acceptance values of the BDT score. The BDT selections have average signal efficiencies of 66.7%, 87.5%, 96.1%, 99.0%, with average background events rejection rates of 95.4%, 31.1%, 27.8% and 11.1%, respectively.

An additional background contribution comes from the process $B \to X_s\gamma$, where $X_s$ is any hadronic state that contains an $s$ quark. To suppress these events, we employ a BDT



JHEP12(2025)109

classifier, $\mathrm{BDT}_{X_s\gamma}$, trained with six variables, separately for the low and high ALP mass regions. These variables include the $E_9/E_{25}$ ratio of each photon, the energy of the most energetic crystal in the ECL cluster, and two helicity angles. The first angle is between the photon momentum and the direction opposite to the $B$ momentum in the $a$ rest frame, and second angle is between the $K^{(*)}$ momentum and the direction opposite to the c.m. system in the $B$ rest frame. The most powerful discriminating variables for all mass hypotheses are the first helicity angle and the energy of the most energetic photon in the c.m. frame. The $\mathrm{BDT}_{X_s\gamma}$ score is not included in the global BDT selection optimization described above. Since this classifier targets a particular specific background process, global optimization may not be appropriate. Instead, we evaluate several threshold values $\mathrm{BDT}_{X_s\gamma}$ and select the one that yields the best PFM for $X_s\gamma$ suppression. The BDT for $X_s\gamma$ suppression achieves an average signal efficiency of 76.6%, with a background rejection rate of 53.0%. Distributions of the BDT classifier scores can be found in appendix B.

The signal-selection efficiency varies across different kaon modes, and within each $K^{(*)}$ mode, the efficiency depends on the ALP mass, which strongly correlates with the $K^{(*)}$ momenta.

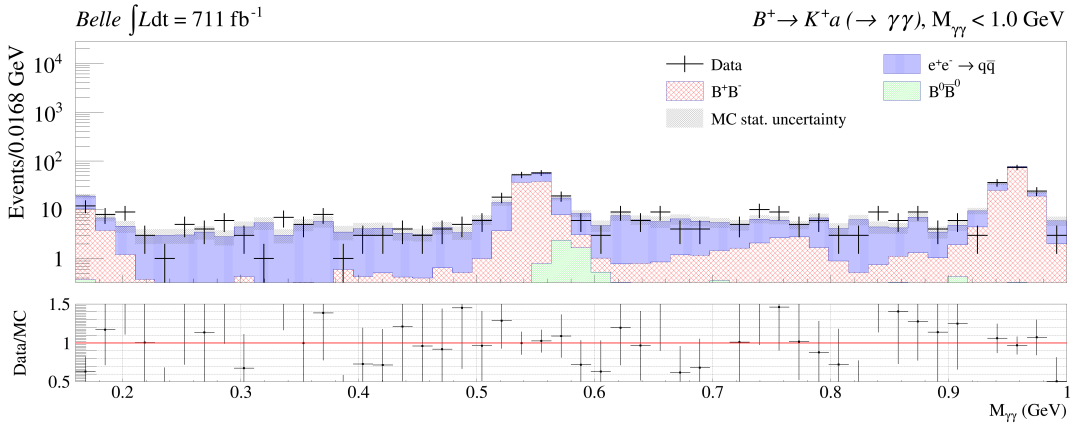
After all selection requirements, the signal selection efficiency varies across different kaon modes. The $K^+$ mode shows the highest efficiency, but it also depends on the ALP mass. In the low-mass region, the efficiency varies between 8% and 10%, while in the high-mass region it shows a stronger variation, reaching a minimum of 7.5% and peaking at about 16% near an ALP mass of 3 GeV. The $K^0_S$ mode follows with slightly lower efficiency between 6% and 11%. The $K^{*0}$ mode shows further reduced efficiency between 3% and 8%. The $K^{*+}$ mode has the lowest efficiency, around 2% across the entire mass range.

## 4 Simulation calibration

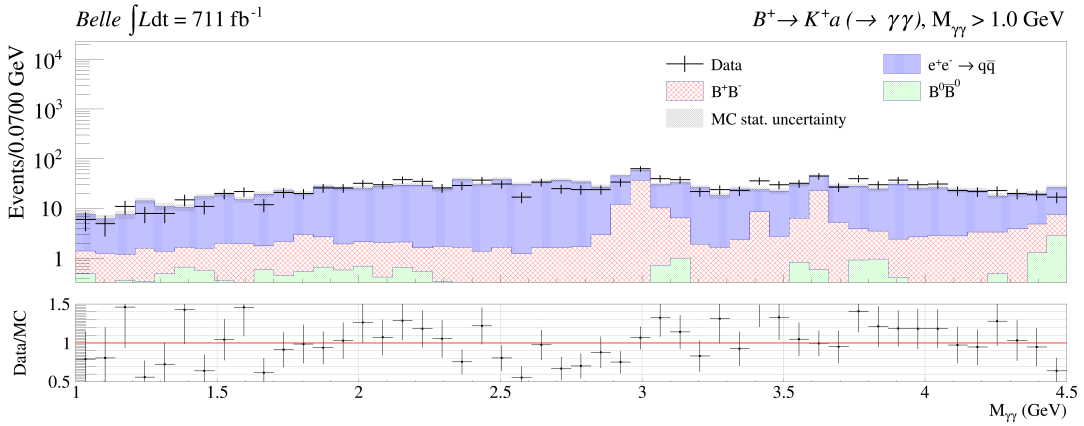
We use the off-resonance data to evaluate corrections to the simulation by comparing experimental data with simulated distributions. The off-resonance data are collected 60 MeV below the $\Upsilon(4S)$ resonance energy, where $e^+e^-$ collisions produce all processes except for $B\overline{B}$ pair production. There are two types of discrepancy found between the experimental data and the simulation: the number of background events, and the shape of the event kinematic and topology variables. For the former, we derive the ratios between the off-resonance data and simulation for each ALP mass range as weights, which are then applied to both the off-resonance and on-resonance continuum simulations. For the latter, we adopt a data-driven method [49]. A BDT classifier is trained to distinguish the differences between the off-resonance data and simulation, using the same variables applied for continuum suppression. Given the BDT output $p(i)$ with $0 < p(i) < 1$ for each simulation event $i$, we calculate the weight $w(i) = p(i)/(1 - p(i))$, which is then applied as a "continuum shape correction" factor. This procedure gives greater importance to simulated events that most closely resemble the real data, improving the agreement between the weighted simulated and experimental distributions. Reweighting based on the off-resonance data enhances the simulated two-photon invariant mass distribution, which then matches the experimental distribution within statistical uncertainty. The resulting on-resonance $M_{\gamma\gamma}$ distributions for low and high ALP mass regions in the $B^+ \to K^+a$ decay are shown in figure 2.



JHEP12(2025)109




(a)




(b)

**Figure 2.** Diphoton invariant mass distribution of ALP candidates in $B^+ \to K^+ a$ decay, overlaid with simulated background contributions from $e^+e^- \to q\bar{q}$ (blue vertical hatched), $e^+e^- \to \Upsilon(4S) \to B^+B^-$ (red cross-hatched), and $e^+e^- \to \Upsilon(4S) \to B^0\overline{B}^0$ (green diagonal hatched) normalized to the experimental data luminosity, with all weights applied.

## 5 Signal extraction and validations

The signal yield $N_{\rm Sig}$ and its standard deviation $\sigma_{\rm Sig}$ are obtained from an unbinned maximum likelihood fit to the $M_{\gamma\gamma}$ distribution. We use a double-sided Crystal Ball function [50] to model the signal $M_{\gamma\gamma}$ distribution. The non peaking background in the $M_{\gamma\gamma}$ distribution is parametrized using a second-order polynomial. The peaking background components from $h \to \gamma\gamma$ ($h = \pi^0$, $\eta$, or $\eta'$) are modelled with double-sided Crystal Ball functions.

We perform a mass scan with a step size equal to the high-side mass resolution parameter from the Crystal Ball signal, $\sigma^R_{\gamma\gamma}$. The latter varies with the ALP mass, ranging from 7.8 MeV at $m_a = 0.160$ GeV to 19.4 MeV at $m_a = 1.9$ GeV, and decreasing to 17.9 MeV at $m_a = 4.5$ GeV. The improved mass resolution in the highest $m_a$ region results from kinematic fitting. Each fit range extends over an $M_{\gamma\gamma}$ interval with a width of $9 \times (\sigma^R_{\gamma\gamma} + \sigma^L_{\gamma\gamma})$, where $\sigma^L_{\gamma\gamma}$ is the low-side Crystal Ball resolution parameter. The signal peak shape parameters depend on the ALP mass, and are derived from the corresponding signal samples. The shape



JHEP12(2025)109

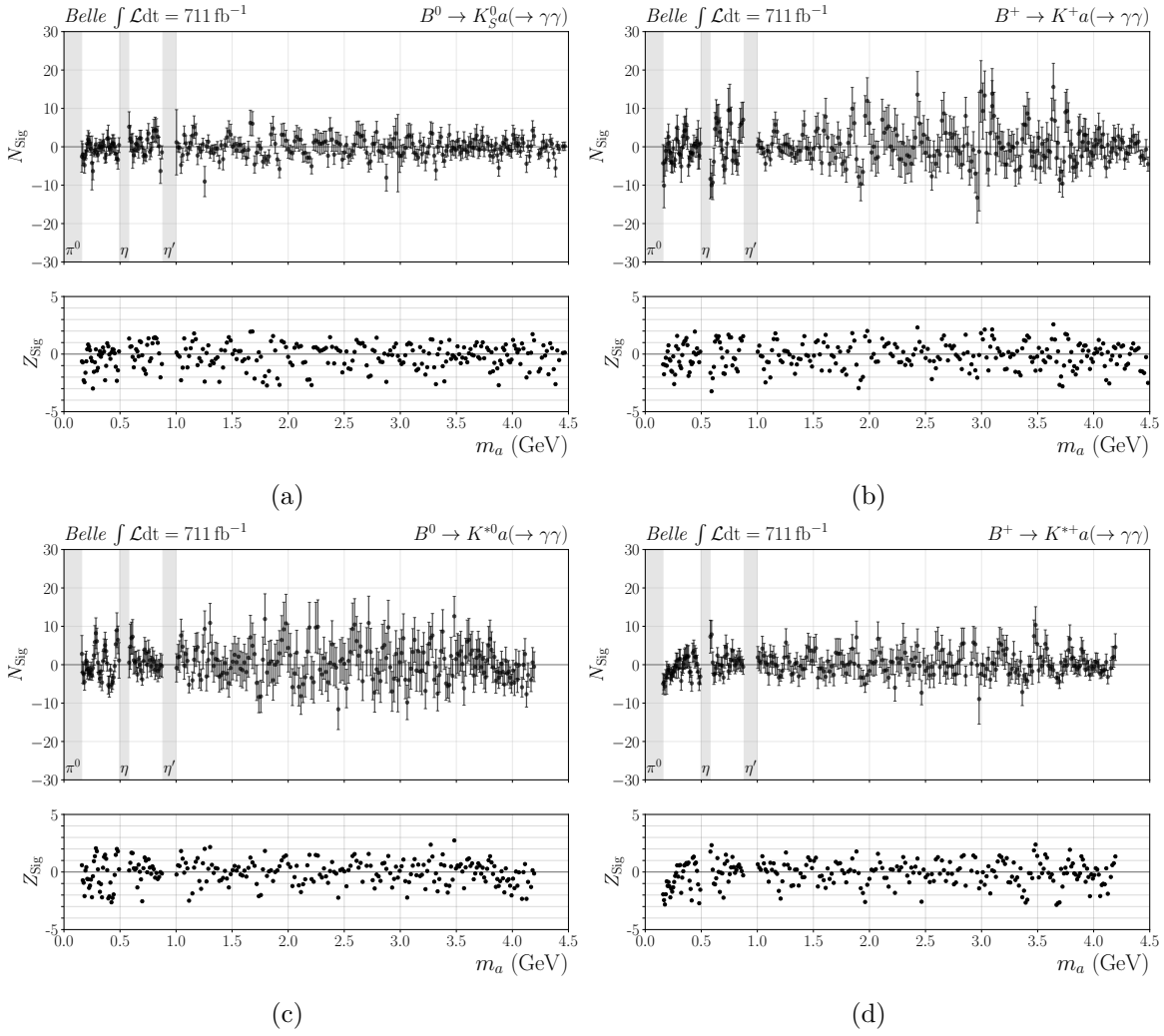



**Figure 3.** Extracted signal yield ($N_{\mathrm{Sig}}$) (top) and the significance level ($Z_{\mathrm{Sig}}$) (bottom) for four kaon modes. The grey bands are the excluded regions corresponding to the $\pi^0, \eta$ and $\eta'$ mass regions.

parameters and position of the peaking background are fixed based on values obtained from the simulation. The combinatorial background parameters are floating in the fit, along with the signal and peaking background normalization, other than the $\eta_c$.

Due to the peaking background from $\pi^0$, $\eta$ and $\eta'$ decays, masses below 0.160 GeV, in the ranges 0.497 GeV–0.578 GeV and 0.938–0.997 GeV, are excluded from the ALP mass scan. The latter two correspond to $\pm 3\sigma$ mass resolution windows centred on the $\eta$ and $\eta'$, respectively. Note that despite these exclusions, the tails of these peaking background components extend into the fitting regions. The ALP mass ranges close to the kinematic limit, from 4.50 GeV to 4.78 GeV for $K$ and from 4.20 GeV to 4.38 GeV for $K^*$, are excluded from our analysis due to low signal efficiency and insufficient background population for reliable signal extraction. The $\eta_c$ mass region is included in the ALP mass scan, as the branching fraction of $B \to K^{(*)}\eta_c(\to \gamma\gamma)$ is sufficiently small that its contribution can be adequately controlled. The normalization factor for the $\Upsilon(4S) \to B\overline{B}$ background component involving $\eta_c$



JHEP12(2025)109

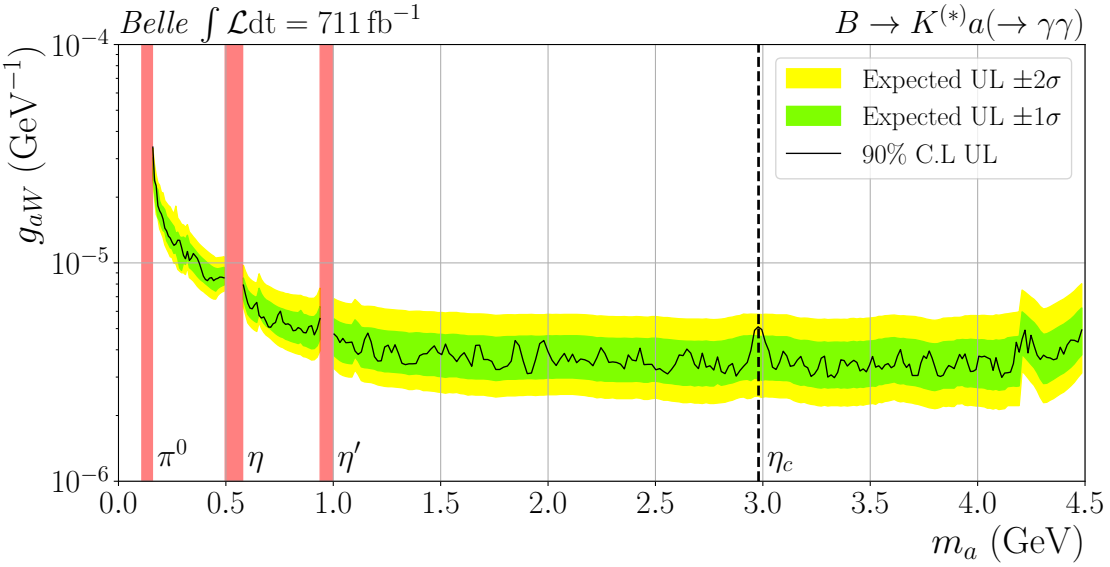



**Figure 4.** 90% CL upper limits on the coupling $g_{aW}$ as a function of the ALP mass obtained with the CLs method with simultaneous fit to the four kaon modes. The green and yellow bands are the $\pm 1$ and $\pm 2$ standard deviation ranges, respectively, for the expected upper limits in the background only model. The red bands are the excluded $\pi^0, \eta$ and $\eta'$ mass regions. The vertical dashed line indicates the nominal $\eta_c$ mass. Systematic uncertainties are included in the figure.

is constrained to the world-average value within the SM uncertainty, while the normalization of the $\eta_c$ component in the $e^+e^- \to q\bar{q}$ background is treated as a free parameter in the fit.

To validate the signal extraction method, we measure the branching fraction of $B \to Kh(\to \gamma\gamma)$ using the fitting procedure described above. The branching fractions of the $B^+ \to K^+\eta$, $B^+ \to K^+\eta'$ and $B^+ \to K^+\eta_c$ decays are measured to be $(0.95 \pm 0.27) \times 10^{-6}$, $(1.79 \pm 0.45) \times 10^{-6}$ and $(2.37 \pm 0.91) \times 10^{-7}$, respectively, consistent with both the previous Belle results [51] and world average values excluding the Belle result [40]. The parametrisation of signal shapes are also validated with these control modes. In all control modes, the shape parameters measured from data and simulated samples are compatible.

In addition, a toy Monte Carlo (ToyMC) study [52] — a simplified, fast simulation in which observables are sampled from probability density functions (p.d.f.) — is carried out to evaluate the fitting bias and signal sensitivity. Ten thousand pseudo-datasets are generated for each ALP mass hypothesis across the four kaon modes using the fitted p.d.f. and the Poisson distribution of signal and background yield obtained from the fit. The result of fits on pseudo-data from the ToyMC has a small negative bias of $4.2\% \times \sigma_{\mathrm{Sig}}$ on average, where $\sigma_{\mathrm{Sig}}$ is the width of the signal yield distributions obtained with the ToyMC study. This is applied as a correction factor to the signal yields in data.

For all ALP mass hypotheses and kaon modes, the decay yield of $B \to K^{(*)}a(\to \gamma\gamma)$ is measured under the assumption that all signal events originate from the prompt decay of the ALP. However, the ALP can be long-lived, and the displaced vertex reduces the signal efficiency. The primary reason for the reduced signal efficiency is the assumption that the photons originate from the IP. The calculated opening angle between the two photons, which



JHEP12(2025)109

enters the ALP mass calculation, is systematically smaller than the true value, which produces a low-side tail on the reconstructed mass distribution. The fit using Crystal Ball parameters from prompt decays systematically underestimates the number of signal events for displaced vertices, effectively reducing the efficiency. To quantify this effect, signal processes with lifetimes $c\tau$ of 10 mm, 50 mm, 100 mm, 200 mm, 300 mm, 400 mm and 500 mm are generated. The decrease in the signal efficiency is modelled as

$$\frac{\varepsilon_{\rm Sig}(c\tau)}{\varepsilon_{\rm Sig}(0)} = re^{a_1 c\tau} + (1-r)e^{a_2 c\tau}, \tag{5.1}$$

where $\varepsilon_{\rm Sig}(0)$ and $\varepsilon_{\rm Sig}(c\tau)$ are the reconstruction efficiencies of prompt and long-lived ALPs, $a_1$, $a_2$ and $r$ are floating parameters obtained from fits to the long-lived ALP simulation results for each ALP mass hypothesis. The resulting functions are incorporated into the upper limit calculation. Consequently, ALP signals with relatively low mass have a longer lifetime and lower signal efficiency, leading to less stringent limits on $g_{aW}$.

The significance is evaluated as $Z_{\rm Sig} = \sqrt{2(L_{\rm s+b} - L_{\rm b})}$, where $L_{\rm s+b}$ and $L_{\rm b}$ are the negative log-likelihoods of the fits with and without signal hypothesis, respectively. The largest observed positive local significance is $2.74\sigma$ at $m_a = 3.482\,\text{GeV}$ in the $K^{*0}$ mode. This local significance corresponds to $1.89\sigma$ global significance, after including the look-elsewhere effect [53]. As shown in figure 3, no significant excess over background is observed and we set 90% confidence level (CL) upper limits on the coupling $g_{aW}$ using the CLs method [54, 55]. For each kaon mode and ALP mass hypothesis, we obtain the branching fractions of $B \to K^{(*)}a(\to \gamma\gamma)$ decays as

$$\mathcal{B}(B \to K^{(*)}a(\to \gamma\gamma)) = \frac{N_{\rm Sig}}{(2 \times N_{\Upsilon(4S)} \times f^x \times \varepsilon_{\rm Sig})}, \tag{5.2}$$

where $N_{\Upsilon(4S)}$ is the total number of $\Upsilon(4S)$ mesons, $f^x$ is the production fraction of $B\overline{B}$ pairs for the neutral mode, $f^{00}$, or the charged mode, $f^{+-}$ [56], and $\varepsilon_{\rm Sig}$ is the signal efficiency. For each ALP mass hypothesis, a simultaneous fit is performed on four kaon modes to obtain $g_{aW}$. Figure 4 shows the resulting limit on the coupling constant $g_{aW}$ as a function of mass. The expected limits obtained with the background only hypothesis are also shown as green and yellow bands. For $m_a > 4.2\,\text{GeV}$, only the $K^+$ and $K^0_S$ modes contribute to the upper limit calculation, as the $K^*$ modes do not extend to this mass range. The apparent behaviour observed near 4.2 GeV reflects this.

## 6 Systematic uncertainties

The systematic uncertainties that affect the extraction of $g_{aW}$ can be classified into two main categories: those originating from prompt decay signal and from long-lived ALP efficiency (table 1).

The uncertainty on the prompt decay signal is derived from our most precise control sample, $B^0 \to K^{*0}\eta$. The obtained branching fraction $\mathcal{B}(B^0 \to K^{*0}\eta) \times \mathcal{B}(\eta \to \gamma\gamma) = (6.8 \pm 1.3) \times 10^{-6}$ is in good agreement with the world-average value of $(6.3 \pm 0.4) \times 10^{-6}$ [40]. We therefore assign a systematic uncertainty of 21%, calculated as the quadratic sum of the fractional branching fraction discrepancy and the associated uncertainties, as a conservative



JHEP12(2025)109

| Source | $K^0_S$ mode | $K^+$ mode | $K^{*0}$ mode | $K^{*+}$ mode |
|---|---|---|---|---|
| Prompt decay signal | | 21.0 | | |
| ⊢ Continuum shape correction [57] | | 4.1 | | |
| ⊢ Photon-detection efficiency [58] | | 4.0 | | |
| ⊢ $K^+$ identification efficiency | — | 3.6 | 3.6 | — |
| ⊢ $f^{00}$ or $f^{+-}$ [56] | 1.7 | 2.1 | 1.7 | 2.1 |
| ⊢ $K^0_S$ reconstruction efficiency [59] | 1.6 | — | — | 1.6 |
| ⊢ $N_{\Upsilon(4S)}$ [60] | | 1.4 | | |
| ⊢ Tracking efficiency [58] | 0.7 | 0.4 | 0.7 | 1.1 |
| Long-lived ALP efficiency | | 7.4 | | |
| Total | | 22.3 | | |

**Table 1.** Summary of the systematic uncertainties (%). The "Prompt decay signal" term is derived from a control sample: it includes the indented items that follow, as well as additional effects (see the text for details).

estimate. The indented items listed under the prompt decay signal in table 1 represent numerically estimable components that affect the prompt decay signal calculation. Other sources not explicitly listed in the table, such as MVA selection efficiency, could not be individually determined and are thus incorporated into the total prompt decay signal alongside the indented items. Consequently, the quadratic sum of the individual components differs from the overall prompt decay signal value.

The systematic uncertainty of the long-lived ALP signal efficiency constraint is derived from $K^0_S \to \pi^0\pi^0$ decays reconstructed in $D^{*+} \to D^0(\to K^0_S\pi^+\pi^-)\pi^+$ events by treating the $K^0_S$ as ALP. Using the same method used for long-lived ALP decays (eq. (5.1)), we construct the signal efficiency function for $K^0_S \to \pi^0\pi^0$ and determine the expected mass resolution on $M(\pi^0\pi^0)$, the invariant mass of the $\pi^0\pi^0$ system, from this function and $K^0_S$ simulation with zero lifetime. The $M(\pi^0\pi^0)$ mass resolution is extracted via p.d.f. fitting in both data and simulation, following the same approach used in ALP signal extraction. The fractional difference between the obtained value and the expectation is $(6.0 \pm 4.4)\%$, which, by adding the difference and its error in quadrature, leads to a conservative estimation of the systematic uncertainty of 7.4% for the long-lived ALP efficiency calculation. The efficiency for long-lived ALP signals decreases as the ALP mass increases, since higher masses correspond to shorter lifetimes. This effect becomes negligibly small for masses above 2.0 GeV. However, as our analysis is dominated by statistical uncertainties, and since mass-dependent refinements to systematic uncertainties would not significantly impact the final results, we have applied the most conservative systematic uncertainty estimate across the entire mass range in our calculations.

The total systematic uncertainty is obtained by summing the prompt decay signal and long-lived ALP efficiency uncertainties in quadrature. The resulting systematic uncertainty is found to be 22.3% for all kaon modes and ALP mass hypotheses. The resulting systematic uncertainty is included in the upper limit calculation procedure by convolving an appropriate Gaussian function with the signal likelihood.



JHEP12(2025)109

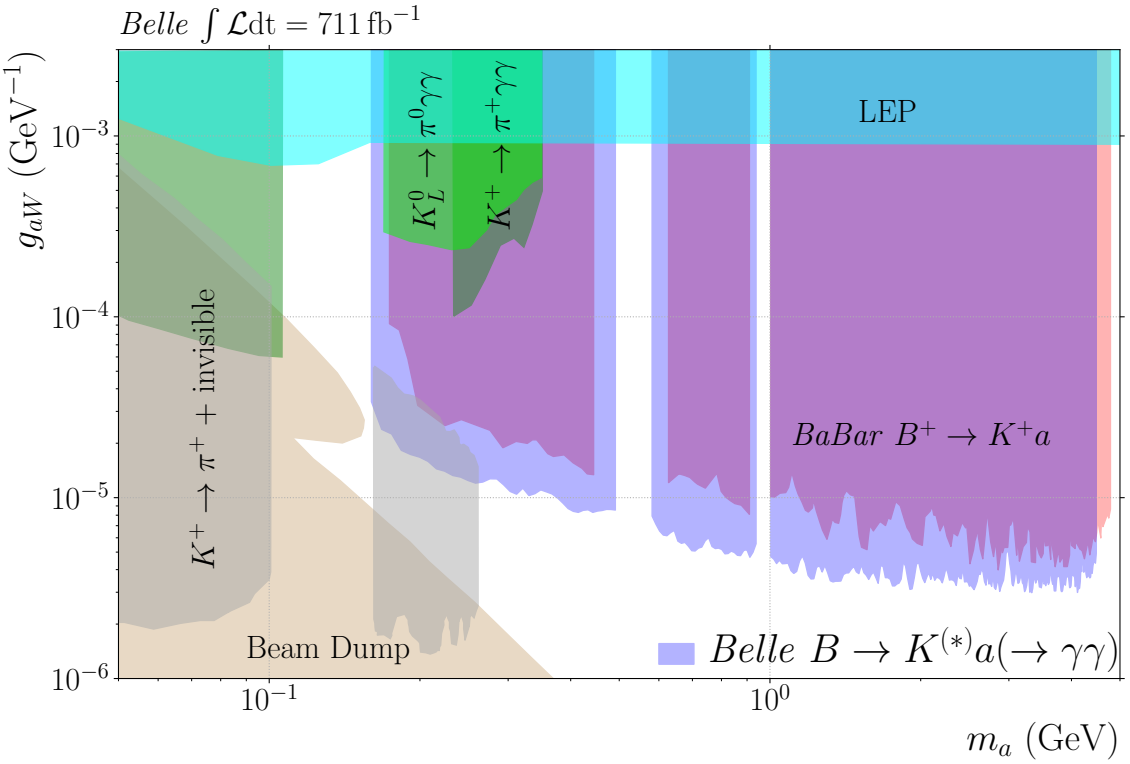



**Figure 5.** The 90% CL upper limits on the coupling $g_{aW}$ from a simultaneous fit to the four $B \to K^{(*)}a$ modes as a function of the ALP mass, compared with existing constraints [28, 29, 61, 62].

## 7 Result and conclusion

We report a search for an axion-like particle in $B \to K^{(*)}a(\to \gamma\gamma)$ decays using a 711 fb$^{-1}$ data sample collected by the Belle experiment at the KEKB $e^+e^-$ collider at a centre-of-mass energy of 10.58 GeV. We search for the decay of the axion-like particle into a pair of photons, $a \to \gamma\gamma$, and explore four kaon modes, $K_S^0$, $K^+$, $K^{*0}$ and $K^{*+}$. We scan the two-photon invariant mass in the range 0.16 GeV–4.50 GeV for the $K$ modes and 0.16 GeV–4.20 GeV for the $K^*$ modes. No significant signal is observed in any of the modes. Figure 5 shows the resulting 90% confidence level upper limits on the $g_{aW}$ coupling as a function of $m_a$, derived from the combination of four kaon modes. The limits are $3 \times 10^{-6}$ GeV$^{-1}$ for the ALP mass hypotheses above 2.0 GeV, increasing to $3 \times 10^{-5}$ GeV$^{-1}$ at the lowest ALP mass. This trend is due to the increase in the lifetime, which leads to a lower signal efficiency. Figure 5 also shows the constraints derived from the NA62 $K^+ \to \pi^+ +$ invisible search [61]. Based on the methodology presented in ref. [62] we reinterpret the NA62 results on a dark scalar decaying to SM particles as limits on ALPs. The constraints on the coupling of the axion-like particle to electroweak gauge bosons $g_{aW}$ are improved by a factor of two compared to the most stringent previous experimental results. [29]



JHEP12(2025)109

## Acknowledgments

This work, based on data collected using the Belle II detector, which was built and commissioned prior to March 2019, was supported by Higher Education and Science Committee of the Republic of Armenia Grant No. 23LCG-1C011; Australian Research Council and Research Grants No. DP200101792, No. DP210101900, No. DP210102831, No. DE220100462, No. LE210100098, and No. LE230100085; Austrian Federal Ministry of Education, Science and Research, Austrian Science Fund (FWF) Grants DOI: 10.55776/P34529, DOI: 10.55776/J4731, DOI: 10.55776/J4625, DOI: 10.55776/M3153, and DOI: 10.55776/PAT1836324, and Horizon 2020 ERC Starting Grant No. 947006 "InterLeptons"; Natural Sciences and Engineering Research Council of Canada, Digital Research Alliance of Canada, and Canada Foundation for Innovation; National Key R&D Program of China under Contract No. 2024YFA1610503, and No. 2024YFA1610504 National Natural Science Foundation of China and Research Grants No. 11575017, No. 11761141009, No. 11705209, No. 11975076, No. 12135005, No. 12150004, No. 12161141008, No. 12405099, No. 12475093, and No. 12175041, and Shandong Provincial Natural Science Foundation Project ZR2022JQ02; the Czech Science Foundation Grant No. 22-18469S, Regional funds of EU/MEYS: OPJAK FORTE CZ.02.01.01/00/22_008/0004632 and Charles University Grant Agency project No. 246122; European Research Council, Seventh Framework PIEF-GA-2013-622527, Horizon 2020 ERC-Advanced Grants No. 267104 and No. 884719, Horizon 2020 ERC-Consolidator Grant No. 819127, Horizon 2020 Marie Sklodowska-Curie Grant Agreement No. 700525 "NIOBE" and No. 101026516, and Horizon 2020 Marie Sklodowska-Curie RISE project JENNIFER2 Grant Agreement No. 822070 (European grants); L'Institut National de Physique Nucléaire et de Physique des Particules (IN2P3) du CNRS and L'Agence Nationale de la Recherche (ANR) under Grant No. ANR-23-CE31-0018 (France); BMFTR, DFG, HGF, MPG, and AvH Foundation (Germany); Department of Atomic Energy under Project Identification No. RTI 4002, Department of Science and Technology, and UPES SEED funding programs No. UPES/R&D-SEED-INFRA/17052023/01 and No. UPES/R&D-SOE/20062022/06 (India); Israel Science Foundation Grant No. 2476/17, U.S.-Israel Binational Science Foundation Grant No. 2016113, and Israel Ministry of Science Grant No. 3-16543; Istituto Nazionale di Fisica Nucleare and the Research Grants BELLE2, and the ICSC – Centro Nazionale di Ricerca in High Performance Computing, Big Data and Quantum Computing, funded by European Union – NextGenerationEU; Japan Society for the Promotion of Science, Grant-in-Aid for Scientific Research Grants No. 16H03968, No. 16H03993, No. 16H06492, No. 16K05323, No. 17H01133, No. 17H05405, No. 18K03621, No. 18H03710, No. 18H05226, No. 19H00682, No. 20H05850, No. 20H05858, No. 22H00144, No. 22K14056, No. 22K21347, No. 23H05433, No. 26220706, and No. 26400255, and the Ministry of Education, Culture, Sports, Science, and Technology (MEXT) of Japan; National Research Foundation (NRF) of Korea Grants No. 2021R1-F1A-1064008, No. 2022R1-A2C-1003993, No. 2022R1-A2C-1092335, No. RS-2016-NR017151, No. RS-2018-NR031074, No. RS-2021-NR060129, No. RS-2023-00208693, No. RS-2024-00354342 and No. RS-2025-02219521, Radiation Science Research Institute, Foreign Large-Size Research Facility Application Supporting project, the Global Science Experimental Data Hub Center, the Korea Institute of Science and Technology Information (K25L2M2C3) and KREONET/GLORIAD; Universiti Malaya RU grant, Akademi Sains Malaysia, and Ministry of Education Malaysia; Frontiers of

JHEP12(2025)109

Science Program Contracts No. FOINS-296, No. CB-221329, No. CB-236394, No. CB-254409, and No. CB-180023, and SEP-CINVESTAV Research Grant No. 237 (Mexico); the Polish Ministry of Science and Higher Education and the National Science Center; the Ministry of Science and Higher Education of the Russian Federation and the HSE University Basic Research Program, Moscow; University of Tabuk Research Grants No. S-0256-1438 and No. S-0280-1439 (Saudi Arabia), and Researchers Supporting Project number (RSPD2025R873), King Saud University, Riyadh, Saudi Arabia; Slovenian Research Agency and Research Grants No. J1-50010 and No. P1-0135; Ikerbasque, Basque Foundation for Science, State Agency for Research of the Spanish Ministry of Science and Innovation through Grant No. PID2022-136510NB-C33, Spain, Agencia Estatal de Investigacion, Spain Grant No. RYC2020-029875-I and Generalitat Valenciana, Spain Grant No. CIDEGENT/2018/020; The Knut and Alice Wallenberg Foundation (Sweden), Contracts No. 2021.0174, No. 2021.0299, and No. 2023.0315; National Science and Technology Council, and Ministry of Education (Taiwan); Thailand Center of Excellence in Physics; TUBITAK ULAKBIM (Turkey); National Research Foundation of Ukraine, Project No. 2020.02/0257, and Ministry of Education and Science of Ukraine; the U.S. National Science Foundation and Research Grants No. PHY-1913789 and No. PHY-2111604, and the U.S. Department of Energy and Research Awards No. DE-AC06-76RLO1830, No. DE-SC0007983, No. DE-SC0009824, No. DE-SC0009973, No. DE-SC0010007, No. DE-SC0010073, No. DE-SC0010118, No. DE-SC0010504, No. DE-SC0011784, No. DE-SC0012704, No. DE-SC0019230, No. DE-SC0021274, No. DE-SC0021616, No. DE-SC0022350, No. DE-SC0023470; and the Vietnam Academy of Science and Technology (VAST) under Grants No. NVCC.05.02/25-25 and No. DL0000.05/26-27.

These acknowledgements are not to be interpreted as an endorsement of any statement made by any of our institutes, funding agencies, governments, or their representatives.

We thank the SuperKEKB team for delivering high-luminosity collisions; the KEK cryogenics group for the efficient operation of the detector solenoid magnet and IBBelle on site; the KEK Computer Research Center for on-site computing support; the NII for SINET6 network support; and the raw-data centers hosted by BNL, DESY, GridKa, IN2P3, INFN, and the University of Victoria.

JHEP12(2025)109

## A Decay width of $B \to K^{(*)}a(\to \gamma\gamma)$ and coupling strength $g_{aW}$

The $B \to K^{(*)}a(\to \gamma\gamma)$ decay width is given by

$$\begin{aligned}\Gamma(B \to Ka) &= \frac{M_B^3}{64\pi}|g_{abs}|^2 \left(1 - \frac{M_K^2}{M_B^2}\right)^2 f_0^2(m_a^2)\lambda_{Ka}^{1/2}, \\ \Gamma(B \to K^*a) &= \frac{M_B^3}{64\pi}|g_{abs}|^2 A_0^2(m_a^2)\lambda_{K^*a}^{3/2},\end{aligned} \tag{A.1}$$

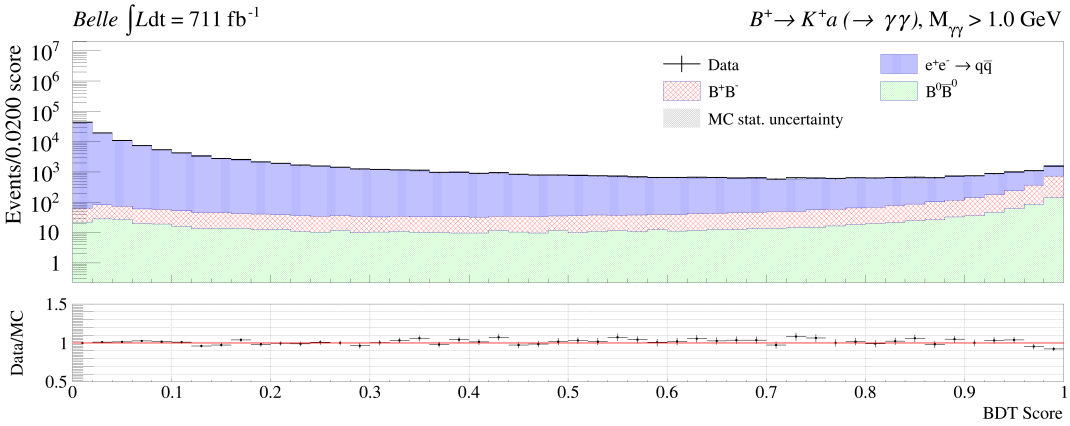



**Figure 6.** Distributions of the first continuum suppression BDT classifier scores from experimental data (black points with error bars) for the $B^+ \to K^+ a(\to \gamma\gamma)$ decay with $m_a > 1.0\,\mathrm{GeV}/c^2$ along with simulated background contributions from $e^+e^- \to q\bar{q}$ (blue vertical hatched), $e^+e^- \to \Upsilon(4S) \to B^+B^-$ (red cross-hatched), and $e^+e^- \to \Upsilon(4S) \to B^0\overline{B}^0$ (green diagonal hatched) normalized to the experimental data luminosity. All correction weights are applied to the continuum events.

where

$$
\begin{aligned}
g_{abs} &= -\frac{3\sqrt{2}G_F M_W^2 g_{aW}}{16\pi^2} \sum_{\alpha=c,t} V_{\alpha b} V_{\alpha s}^* f\left(\frac{M_\alpha^2}{M_W^2}\right), \\
f(x) &\equiv \frac{x[1+x(\log x - 1)]}{(1-x)^2}, \\
f_0(m_a^2) &= \frac{0.330}{1 - m_a^2/37.46}, \\
A_0(m_a^2) &= \frac{1.364}{1 - m_a^2/27.88} - \frac{0.990}{1 - m_a^2/36.78}, \\
\lambda_{K^{(*)}a} &= \left(1 - \frac{(m_a + M_{K^{(*)}})^2}{M_B^2}\right)\left(1 - \frac{(m_a - M_{K^{(*)}})^2}{M_B^2}\right).
\end{aligned}
\tag{A.2}
$$

Here, $M_B$, $M_K$ and $M_W$ are the masses of $B$ meson, kaon and $W$ boson, respectively, while $m_a$ is the ALP mass. The functions $f_0(q)$ and $A_0(q)$ are the form factors from the hadronic matrix elements [63, 64], $G_F$ is the Fermi constant, and $V_{\alpha b}$ and $V_{\alpha s}$ with $\alpha = c, t$ are the corresponding Cabibbo-Kobayashi-Maskawa (CKM) matrix elements.

## B Figures

The following includes additional figures.



JHEP12(2025)109

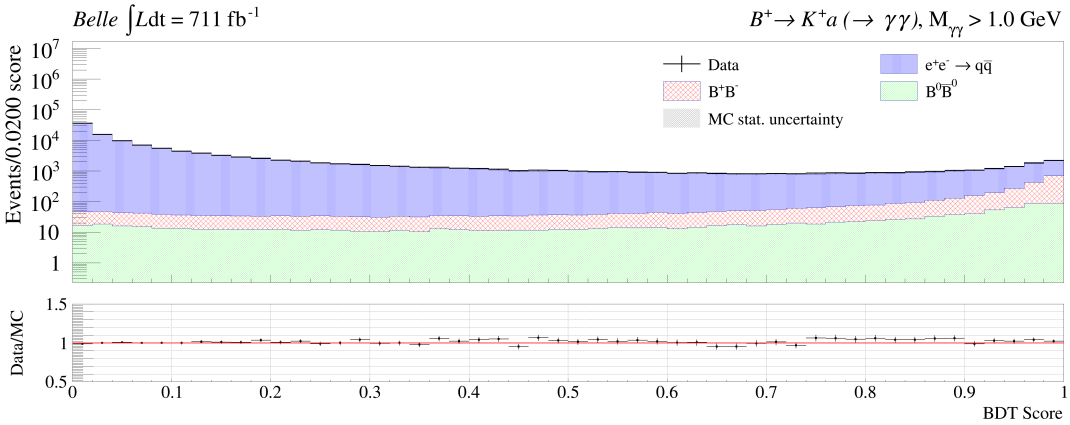



**Figure 7.** Distributions of the second continuum suppression BDT classifier scores. The colour convention used in this histogram is identical to that in figure 6.

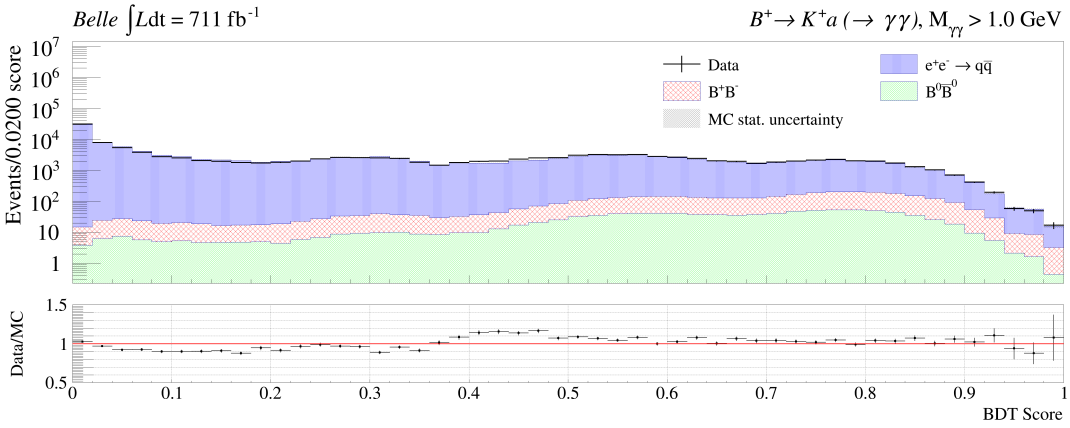



**Figure 8.** Distributions of $P_{\pi^0}(\gamma_1)$ BDT classifier scores. The colour convention used in this histogram is identical to that in figure 6.

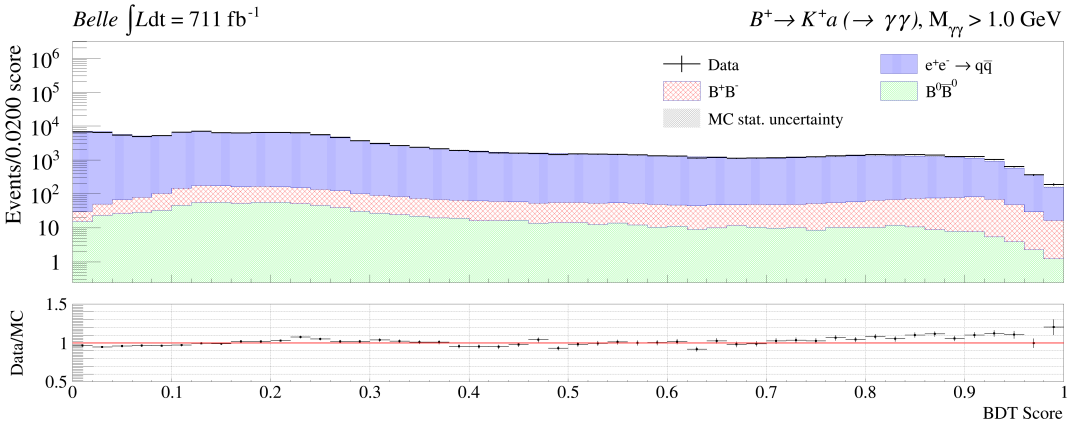



**Figure 9.** Distributions of $P_{\pi^0}(\gamma_2)$BDT classifier scores. The colour convention used in this histogram is identical to that in figure 6.



JHEP12(2025)109

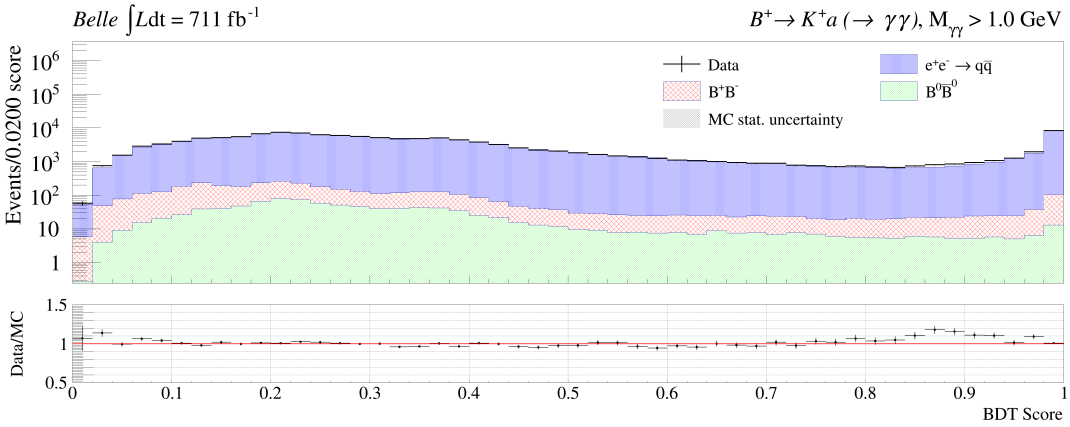
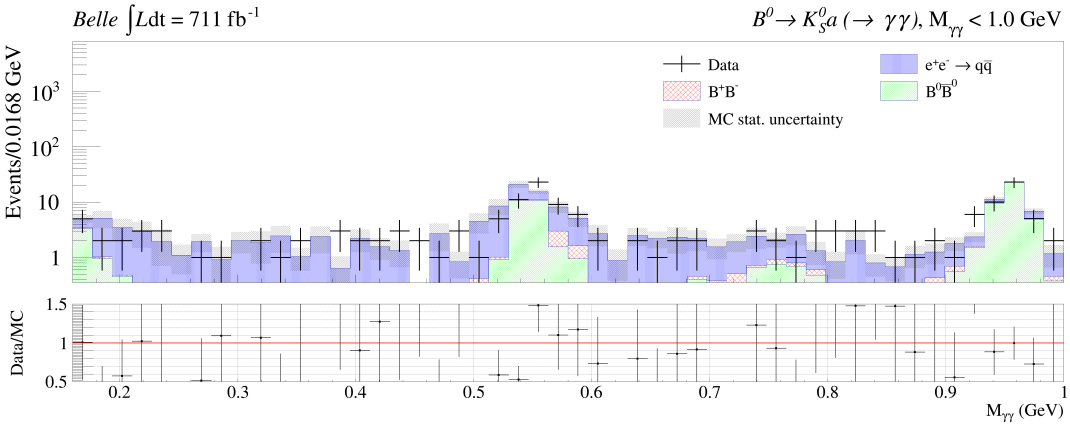
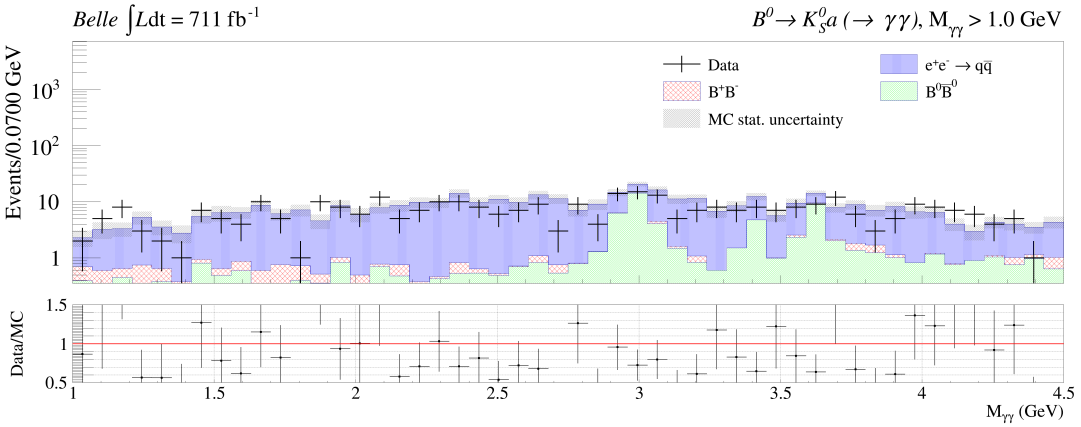



**Figure 10.** Distributions of $X_s\gamma$ suppression BDT classifier scores. The colour convention used in this histogram is identical to that in figure 6.




(a)

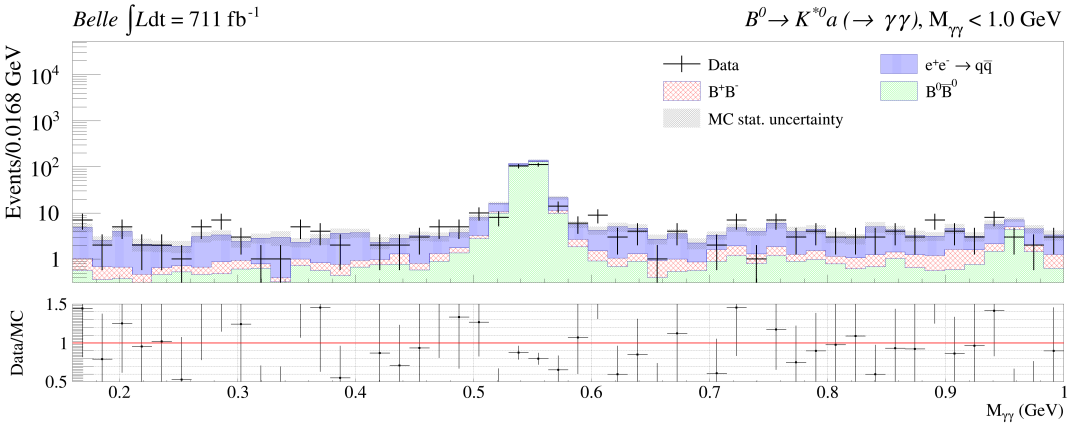



(b)

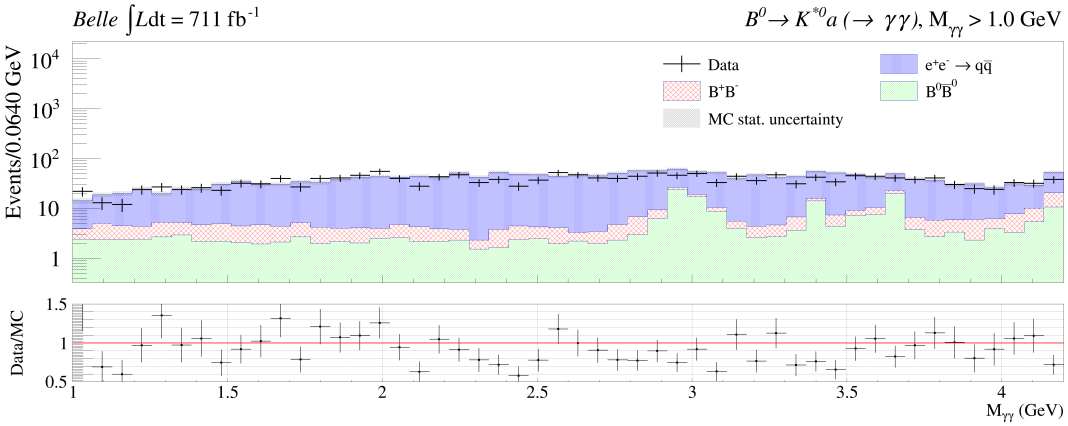
**Figure 11.** Diphoton invariant mass distribution of ALP candidates in $B^0 \to K^0_S a$ decay, The colour convention used in this histogram is identical to that in figure 2.



JHEP12(2025)109



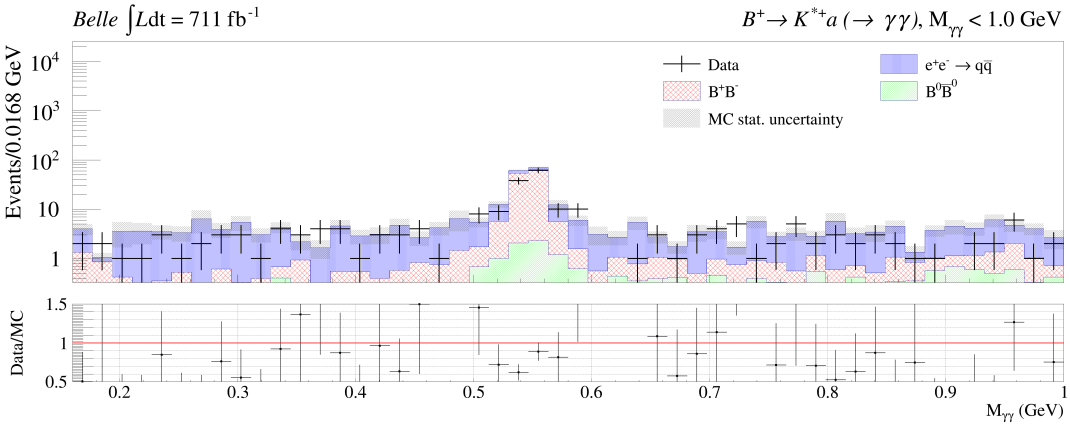

(a)



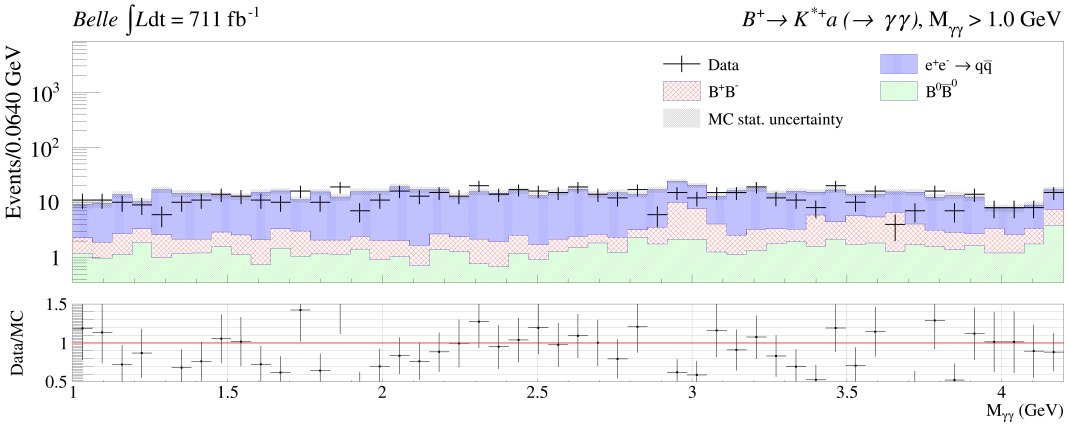

(b)

**Figure 12.** Diphoton invariant mass distribution of ALP candidates in $B^0 \to K^{*0}a$ decay, The colour convention used in this histogram is identical to that in figure 2.

JHEP12(2025)109




(a)




(b)

**Figure 13.** Diphoton invariant mass distribution of ALP candidates in $B^+ \to K^{*+}a$ decay, The colour convention used in this histogram is identical to that in figure 2.

JHEP12(2025)109

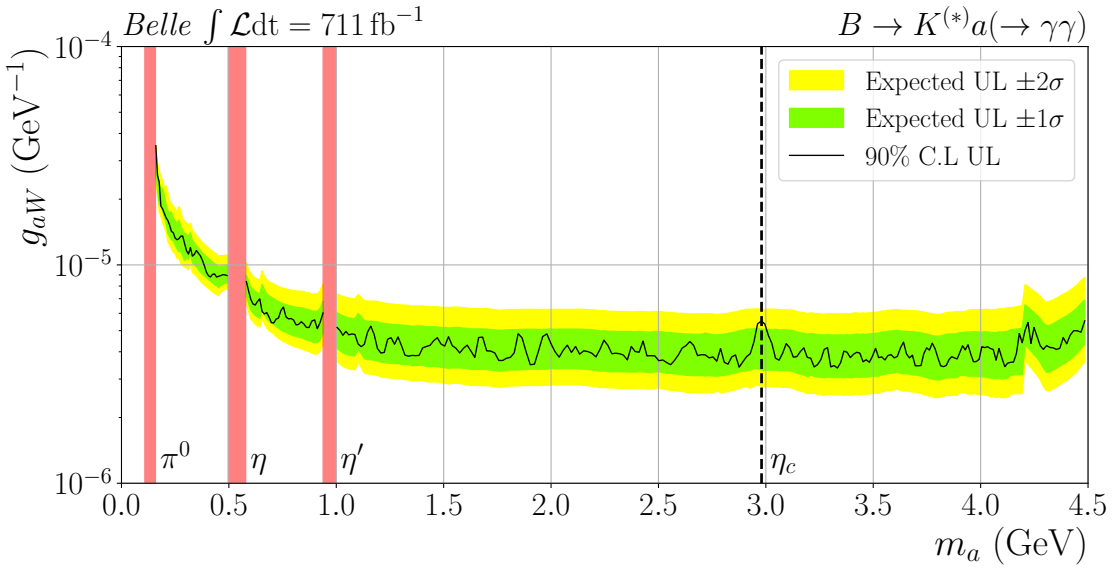



**Figure 14.** The 95% CL upper limits on the coupling $g_{aW}$ as a function of the ALP mass obtained with CLs method with simultaneous fit. The green and yellow bands are $\pm 1$ and $\pm 2$ standard deviation ranges, respectively, of expected upper limit of background only model. The red bands are the excluded $\pi^0, \eta$ and $\eta'$ mass regions. The vertical dashed line indicates the nominal $\eta_c$ mass.

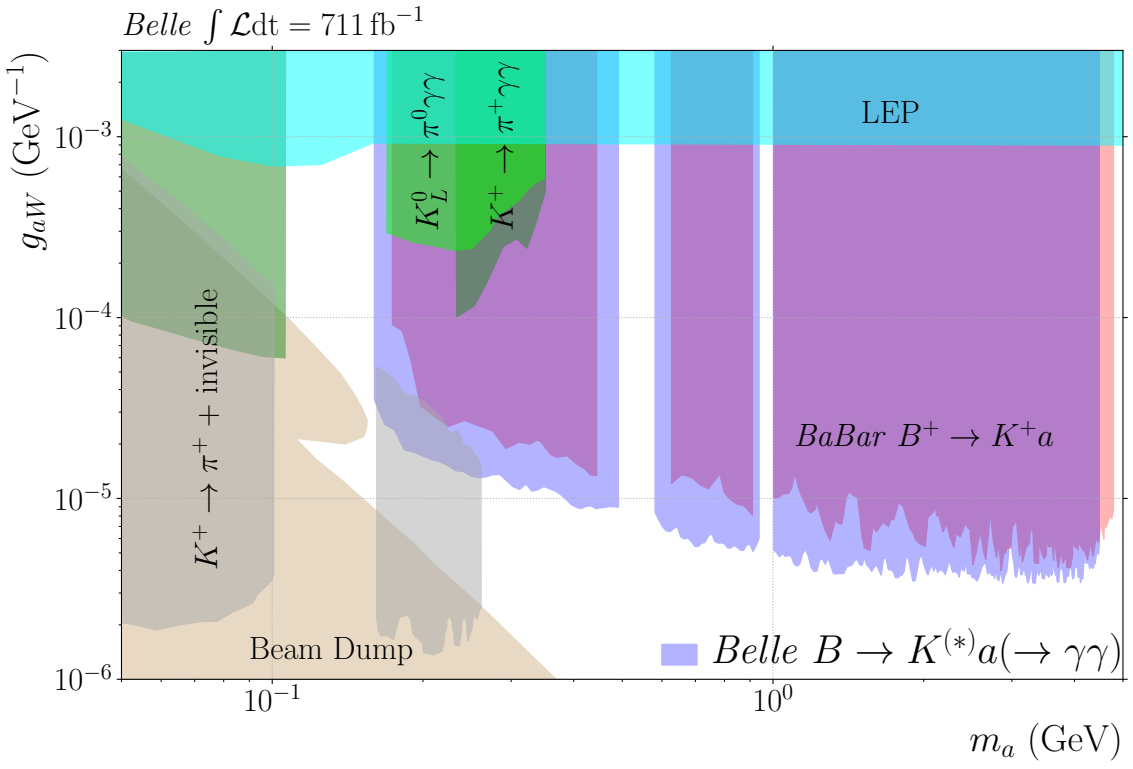



**Figure 15.** The 95% CL upper limits on the coupling $g_{aW}$ from a simultaneous fit to the four $B \to K^{(*)}a$ modes as a function of the ALP mass, compared with existing constraints.

JHEP12(2025)109

**Data Availability Statement.** This article has no associated data or the data will not be deposited.

**Code Availability Statement.** This article has no associated code or the code will not be deposited.



## References

[1] R.D. Peccei and H.R. Quinn, *CP conservation in the presence of instantons*, *Phys. Rev. Lett.* **38** (1977) 1440 [INSPIRE].

[2] S. Weinberg, *A new light boson?*, *Phys. Rev. Lett.* **40** (1978) 223 [INSPIRE].

[3] F. Wilczek, *Problem of strong P and T invariance in the presence of instantons*, *Phys. Rev. Lett.* **40** (1978) 279 [INSPIRE].

[4] J. Preskill, M.B. Wise and F. Wilczek, *Cosmology of the invisible axion*, *Phys. Lett. B* **120** (1983) 127 [INSPIRE].

[5] L.F. Abbott and P. Sikivie, *A cosmological bound on the invisible axion*, *Phys. Lett. B* **120** (1983) 133 [INSPIRE].

[6] M. Dine and W. Fischler, *The not so harmless axion*, *Phys. Lett. B* **120** (1983) 137 [INSPIRE].

[7] Y. Nomura and J. Thaler, *Dark matter through the axion portal*, *Phys. Rev. D* **79** (2009) 075008 [arXiv:0810.5397] [INSPIRE].

[8] K. Mimasu and V. Sanz, *ALPs at colliders*, *JHEP* **06** (2015) 173 [arXiv:1409.4792] [INSPIRE].

[9] M.J. Dolan, F. Kahlhoefer, C. McCabe and K. Schmidt-Hoberg, *A taste of dark matter: flavour constraints on pseudoscalar mediators*, *JHEP* **03** (2015) 171 [*Erratum ibid.* **07** (2015) 103] [arXiv:1412.5174] [INSPIRE].

[10] J. Jaeckel and M. Spannowsky, *Probing MeV to 90 GeV axion-like particles with LEP and LHC*, *Phys. Lett. B* **753** (2016) 482 [arXiv:1509.00476] [INSPIRE].

[11] B. Döbrich et al., *ALPtraum: ALP production in proton beam dump experiments*, *JHEP* **02** (2016) 018 [arXiv:1512.03069] [INSPIRE].

[12] S. Knapen, T. Lin, H.K. Lou and T. Melia, *Searching for axionlike particles with ultraperipheral heavy-ion collisions*, *Phys. Rev. Lett.* **118** (2017) 171801 [arXiv:1607.06083] [INSPIRE].

[13] M.J. Dolan et al., *Revised constraints and Belle II sensitivity for visible and invisible axion-like particles*, *JHEP* **12** (2017) 094 [*Erratum ibid.* **03** (2021) 190] [arXiv:1709.00009] [INSPIRE].

[14] I. Brivio et al., *ALPs effective field theory and collider signatures*, *Eur. Phys. J. C* **77** (2017) 572 [arXiv:1701.05379] [INSPIRE].

[15] M. Bauer, M. Neubert and A. Thamm, *LHC as an axion factory: probing an axion explanation for $(g-2)_\mu$ with exotic Higgs decays*, *Phys. Rev. Lett.* **119** (2017) 031802 [arXiv:1704.08207] [INSPIRE].

[16] K. Choi, S.H. Im, C.B. Park and S. Yun, *Minimal flavor violation with axion-like particles*, *JHEP* **11** (2017) 070 [arXiv:1708.00021] [INSPIRE].



JHEP12(2025)109

[17] M. Bauer, M. Neubert and A. Thamm, *Collider probes of axion-like particles*, *JHEP* **12** (2017) 044 [arXiv:1708.00443] [INSPIRE].

[18] A. Mariotti, D. Redigolo, F. Sala and K. Tobioka, *New LHC bound on low-mass diphoton resonances*, *Phys. Lett. B* **783** (2018) 13 [arXiv:1710.01743] [INSPIRE].

[19] X. Cid Vidal et al., *New axion searches at flavor factories*, *JHEP* **01** (2019) 113 [*Erratum ibid.* **06** (2020) 141] [arXiv:1810.09452] [INSPIRE].

[20] CHARM collaboration, *Search for axion like particle production in 400 GeV proton-copper interactions*, *Phys. Lett. B* **157** (1985) 458 [INSPIRE].

[21] E.M. Riordan et al., *A search for short lived axions in an electron beam dump experiment*, *Phys. Rev. Lett.* **59** (1987) 755 [INSPIRE].

[22] J.D. Bjorken et al., *Search for neutral metastable penetrating particles produced in the SLAC beam dump*, *Phys. Rev. D* **38** (1988) 3375 [INSPIRE].

[23] J. Blumlein et al., *Limits on neutral light scalar and pseudoscalar particles in a proton beam dump experiment*, *Z. Phys. C* **51** (1991) 341 [INSPIRE].

[24] OPAL collaboration, *Multiphoton production in $e^+e^-$ collisions at $\sqrt{s} = 181$ GeV to 209 GeV*, *Eur. Phys. J. C* **26** (2003) 331 [hep-ex/0210016] [INSPIRE].

[25] D. Aloni, Y. Soreq and M. Williams, *Coupling QCD-scale axionlike particles to gluons*, *Phys. Rev. Lett.* **123** (2019) 031803 [arXiv:1811.03474] [INSPIRE].

[26] Belle-II collaboration, *Search for axion-like particles produced in $e^+e^-$ collisions at Belle II*, *Phys. Rev. Lett.* **125** (2020) 161806 [arXiv:2007.13071] [INSPIRE].

[27] LHCb collaboration, *Search for resonances decaying to photon pairs with masses between 4.9 and 19.4 GeV*, arXiv:2507.14390 [INSPIRE].

[28] E. Izaguirre, T. Lin and B. Shuve, *Searching for axionlike particles in flavor-changing neutral current processes*, *Phys. Rev. Lett.* **118** (2017) 111802 [arXiv:1611.09355] [INSPIRE].

[29] BaBar collaboration, *Search for an axionlike particle in B meson decays*, *Phys. Rev. Lett.* **128** (2022) 131802 [arXiv:2111.01800] [INSPIRE].

[30] Belle collaboration, *The Belle detector*, *Nucl. Instrum. Meth. A* **479** (2002) 117 [INSPIRE].

[31] S. Kurokawa and E. Kikutani, *Overview of the KEKB accelerators*, *Nucl. Instrum. Meth. A* **499** (2003) 1 [INSPIRE].

[32] D.J. Lange, *The EvtGen particle decay simulation package*, *Nucl. Instrum. Meth. A* **462** (2001) 152 [INSPIRE].

[33] T. Sjöstrand et al., *An introduction to PYTHIA 8.2*, *Comput. Phys. Commun.* **191** (2015) 159 [arXiv:1410.3012] [INSPIRE].

[34] E. Barberio, B. van Eijk and Z. Was, *PHOTOS: a universal Monte Carlo for QED radiative corrections in decays*, *Comput. Phys. Commun.* **66** (1991) 115 [INSPIRE].

[35] R. Brun, F. Bruyant, M. Maire, A.C. McPherson and P. Zanarini, *GEANT 3: user's guide Geant 3.10, Geant 3.11; rev. version*, CERN, Geneva, Switzerland (1987).

[36] M. Gelb et al., *B2BII: data conversion from Belle to Belle II*, *Comput. Softw. Big Sci.* **2** (2018) 9 [arXiv:1810.00019] [INSPIRE].

[37] Belle-II Framework Software Group collaboration, *The Belle II core software*, *Comput. Softw. Big Sci.* **3** (2019) 1 [arXiv:1809.04299] [INSPIRE].



JHEP12(2025)109

[38] Belle II collaboration, *Belle II analysis software framework (BASF2)*, Zenodo (2025).

[39] Belle collaboration, *Measurement of branching fractions of $\Lambda_c^+ \to pK_S^0K_S^0$ and $\Lambda_c^+ \to pK_S^0\eta$ at Belle*, *Phys. Rev. D* **107** (2023) 032004 [arXiv:2210.01995] [INSPIRE].

[40] Particle Data Group collaboration, *Review of particle physics*, *Phys. Rev. D* **110** (2024) 030001 [INSPIRE].

[41] H. Nakano, *Search for new physics by a time-dependent CP violation analysis of the decay $B \to K_S\eta\gamma$ using the Belle detector*, Ph.D. thesis, Tohoku U., Sendai, Japan (2015) [INSPIRE].

[42] T. Keck, *FastBDT: a speed-optimized multivariate classification algorithm for the Belle II experiment*, *Comput. Softw. Big Sci.* **1** (2017) 2 [INSPIRE].

[43] G.C. Fox and S. Wolfram, *Observables for the analysis of event shapes in $e^+e^-$ annihilation and other processes*, *Phys. Rev. Lett.* **41** (1978) 1581 [INSPIRE].

[44] Belle collaboration, *Evidence for $B^0 \to \pi^0\pi^0$*, *Phys. Rev. Lett.* **91** (2003) 261801 [hep-ex/0308040] [INSPIRE].

[45] Belle collaboration, *Observation of Cabibbo suppressed $B \to D^{(*)}K^-$ decays at BELLE*, *Phys. Rev. Lett.* **87** (2001) 111801 [hep-ex/0104051] [INSPIRE].

[46] BaBar and Belle collaborations, *The physics of the B factories*, *Eur. Phys. J. C* **74** (2014) 3026 [arXiv:1406.6311] [INSPIRE].

[47] J.D. Bjorken and S.J. Brodsky, *Statistical model for electron-positron annihilation into hadrons*, *Phys. Rev. D* **1** (1970) 1416 [INSPIRE].

[48] G. Punzi, *Sensitivity of searches for new signals and its optimization*, *eConf C* **030908** (2003) MODT002 [physics/0308063] [INSPIRE].

[49] D. Martschei, M. Feindt, S. Honc and J. Wagner-Kuhr, *Advanced event reweighting using multivariate analysis*, *J. Phys. Conf. Ser.* **368** (2012) 012028 [INSPIRE].

[50] J. Gaiser, *Charmonium spectroscopy from radiative decays of the $J/\psi$ and $\psi'$*, Ph.D. thesis, Stanford University, Stanford, CA, U.S.A. (1982) [INSPIRE].

[51] Belle collaboration, *Search for resonant $B^\pm \to K^\pm h \to K^\pm\gamma\gamma$ decays at Belle*, *Phys. Lett. B* **662** (2008) 323 [hep-ex/0608037] [INSPIRE].

[52] W. Verkerke and D.P. Kirkby, *The RooFit toolkit for data modeling*, *eConf C* **0303241** (2003) MOLT007 [physics/0306116] [INSPIRE].

[53] E. Gross and O. Vitells, *Trial factors for the look elsewhere effect in high energy physics*, *Eur. Phys. J. C* **70** (2010) 525 [arXiv:1005.1891] [INSPIRE].

[54] G. Cowan, K. Cranmer, E. Gross and O. Vitells, *Asymptotic formulae for likelihood-based tests of new physics*, *Eur. Phys. J. C* **71** (2011) 1554 [*Erratum ibid.* **73** (2013) 2501] [arXiv:1007.1727] [INSPIRE].

[55] A.L. Read, *Presentation of search results: the $CL_s$ technique*, *J. Phys. G* **28** (2002) 2693 [INSPIRE].

[56] Heavy Flavor Averaging Group (HFLAV) collaboration, *Averages of b-hadron, c-hadron, and $\tau$-lepton properties as of 2023*, arXiv:2411.18639 [INSPIRE].

[57] Belle collaboration, *Search for $B^+ \to \mu^+\nu_\mu$ and $B^+ \to \mu^+ N$ with inclusive tagging*, *Phys. Rev. D* **101** (2020) 032007 [arXiv:1911.03186] [INSPIRE].

[58] Belle collaboration, *Search for the radiative penguin decays $B^0 \to K_S^0K_S^0\gamma$ in the Belle experiment*, *Phys. Rev. D* **106** (2022) 012006 [arXiv:2203.05320] [INSPIRE].

JHEP12(2025)109

[59] N. Dash et al., *Search for CP violation and measurement of the branching fraction in the decay* $D^0 \to K^0_S K^0_S$, *Phys. Rev. Lett.* **119** (2017) 171801 [arXiv:1705.05966] [INSPIRE].

[60] BELLE collaboration, *Study of exclusive* $B \to X_u \ell \nu$ *decays and extraction of* $\|V_{ub}\|$ *using full reconstruction tagging at the Belle experiment*, *Phys. Rev. D* **88** (2013) 032005 [arXiv:1306.2781] [INSPIRE].

[61] NA62 collaboration, *Measurement of the very rare* $K^+ \to \pi^+ \nu\overline{\nu}$ *decay*, *JHEP* **06** (2021) 093 [arXiv:2103.15389] [INSPIRE].

[62] E. Goudzovski et al., *New physics searches at kaon and hyperon factories*, *Rept. Prog. Phys.* **86** (2023) 016201 [arXiv:2201.07805] [INSPIRE].

[63] P. Ball and R. Zwicky, $B_{d,s} \to \rho, \omega, K^*, \phi$ *decay form-factors from light-cone sum rules revisited*, *Phys. Rev. D* **71** (2005) 014029 [hep-ph/0412079] [INSPIRE].

[64] P. Ball and R. Zwicky, *New results on* $B \to \pi, K, \eta$ *decay formfactors from light-cone sum rules*, *Phys. Rev. D* **71** (2005) 014015 [hep-ph/0406232] [INSPIRE].

JHEP12(2025)109

## The Belle and Belle II collaborations

I. Adachi, L. Aggarwal, H. Ahmed, Y. Ahn, H. Aihara, N. Akopov, S. Alghamdi, M. Alhakami, A. Aloisio, N. Althubiti, K. Amos, M. Angelsmark, N. Anh Ky, C. Antonioli, D. M. Asner, H. Atmacan, T. Aushev, V. Aushev, M. Aversano, R. Ayad, V. Babu, H. Bae, N. K. Baghel, S. Bahinipati, P. Bambade, Sw. Banerjee, S. Bansal, M. Barrett, M. Bartl, J. Baudot, A. Baur, A. Beaubien, F. Becherer, J. Becker, J. V. Bennett, F. U. Bernlochner, V. Bertacchi, M. Bertemes, E. Bertholet, M. Bessner, S. Bettarini, B. Bhuyan, F. Bianchi, T. Bilka, D. Biswas, A. Bobrov, D. Bodrov, A. Bondar, G. Bonvicini, J. Borah, A. Boschetti, A. Bozek, M. Bračko, P. Branchini, T. E. Browder, A. Budano, S. Bussino, Q. Campagna, M. Campajola, L. Cao, G. Casarosa, C. Cecchi, M.-C. Chang, P. Cheema, L. Chen, B. G. Cheon, K. Chilikin, J. Chin, K. Chirapatpimol, H.-E. Cho, K. Cho, S.-J. Cho, S.-K. Choi, S. Choudhury, J. Cochran, I. Consigny, L. Corona, J. X. Cui, E. De La Cruz-Burelo, S. A. De La Motte, G. De Nardo, G. De Pietro, R. de Sangro, M. Destefanis, S. Dey, A. Di Canto, J. Dingfelder, Z. Doležal, I. Domínguez Jiménez, T. V. Dong, K. Dugic, G. Dujany, P. Ecker, J. Eppelt, R. Farkas, P. Feichtinger, T. Ferber, T. Fillinger, C. Finck, G. Finocchiaro, A. Fodor, F. Forti, A. Frey, B. G. Fulsom, A. Gabrielli, A. Gale, M. Garcia-Hernandez, R. Garg, G. Gaudino, V. Gaur, V. Gautam, A. Gaz, A. Gellrich, G. Ghevondyan, D. Ghosh, H. Ghumaryan, G. Giakoustidis, R. Giordano, A. Giri, P. Gironella Gironell, B. Gobbo, R. Godang, O. Gogota, P. Goldenzweig, E. Graziani, D. Greenwald, Z. Gruberová, Y. Guan, K. Gudkova, I. Haide, Y. Han, T. Hara, K. Hayasaka, H. Hayashii, S. Hazra, C. Hearty, M. T. Hedges, G. Heine, I. Heredia de la Cruz, M. Hernández Villanueva, T. Higuchi, M. Hoek, M. Hohmann, R. Hoppe, P. Horak, C.-L. Hsu, T. Humair, T. Iijima, K. Inami, N. Ipsita, A. Ishikawa, R. Itoh, M. Iwasaki, P. Jackson, W. W. Jacobs, E.-J. Jang, Q. P. Ji, S. Jia, Y. Jin, A. Johnson, K. K. Joo, H. Junkerkalefeld, J. Kandra, K. H. Kang, G. Karyan, T. Kawasaki, F. Keil, C. Ketter, M. Khan, C. Kiesling, C. Kim, C.-H. Kim, D. Y. Kim, J.-Y. Kim, K.-H. Kim, Y. J. Kim, Y.-K. Kim, H. Kindo, K. Kinoshita, P. Kodyš, T. Koga, S. Kohani, K. Kojima, A. Korobov, S. Korpar, E. Kovalenko, P. Križan, P. Krokovny, T. Kuhr, Y. Kulii, D. Kumar, R. Kumar, K. Kumara, T. Kunigo, A. Kuzmin, Y.-J. Kwon, S. Lacaprara, K. Lalwani, T. Lam, J. S. Lange, T. S. Lau, M. Laurenza, R. Leboucher, F. R. Le Diberder, M. J. Lee, C. Lemettais, P. Leo, P. M. Lewis, C. Li, H.-J. Li, L. K. Li, Q. M. Li, W. Z. Li, Y. Li, Y. B. Li, Y. P. Liao, J. Libby, J. Lin, S. Lin, M. H. Liu, Q. Y. Liu, Y. Liu, Z. Q. Liu, D. Liventsev, S. Longo, T. Lueck, C. Lyu, Y. Ma, C. Madaan, M. Maggiora, S. P. Maharana, R. Maiti, G. Mancinelli, R. Manfredi, E. Manoni, M. Mantovano, D. Marcantonio, S. Marcello, C. Marinas, C. Martellini, A. Martens, T. Martinov, L. Massaccesi, M. Masuda, D. Matvienko, S. K. Maurya, M. Maushart, J. A. McKenna, R. Mehta, F. Meier, D. Meleshko, M. Merola, C. Miller, M. Mirra, S. Mitra, K. Miyabayashi, H. Miyake, R. Mizuk, G. B. Mohanty, S. Mondal, S. Moneta, A. L. Moreira de Carvalho, H.-G. Moser, R. Mussa, I. Nakamura, M. Nakao, Y. Nakazawa, M. Naruki, Z. Natkaniec, A. Natochii, M. Nayak, M. Neu, M. Niiyama, S. Nishida, S. Ogawa, H. Ono, G. Pakhlova, S. Pardi, K. Parham, H. Park, J. Park,

JHEP12(2025)109

K. Park, S.-H. Park, B. Paschen, A. Passeri, S. Patra, S. Paul, T. K. Pedlar, I. Peruzzi, R. Peschke, R. Pestotnik, M. Piccolo, L. E. Piilonen, P. L. M. Podesta-Lerma, T. Podobnik, S. Pokharel, A. Prakash, C. Praz, S. Prell, E. Prencipe, M. T. Prim, S. Privalov, H. Purwar, P. Rados, G. Raeuber, S. Raiz, K. Ravindran, J. U. Rehman, M. Reif, S. Reiter, M. Remnev, L. Reuter, D. Ricalde Herrmann, I. Ripp-Baudot, G. Rizzo, S. H. Robertson, J. M. Roney, A. Rostomyan, N. Rout, L. Salutari, D. A. Sanders, S. Sandilya, L. Santelj, V. Savinov, B. Scavino, J. Schmitz, S. Schneider, G. Schnell, M. Schnepf, C. Schwanda, Y. Seino, A. Selce, K. Senyo, J. Serrano, M. E. Sevior, C. Sfienti, W. Shan, G. Sharma, C. P. Shen, X. D. Shi, T. Shillington, T. Shimasaki, J.-G. Shiu, D. Shtol, A. Sibidanov, F. Simon, J. B. Singh, J. Skorupa, M. Sobotzik, A. Soffer, A. Sokolov, E. Solovieva, S. Spataro, B. Spruck, M. Starič, P. Stavroulakis, S. Stefkova, R. Stroili, Y. Sue, M. Sumihama, K. Sumisawa, N. Suwonjandee, H. Svidras, M. Takizawa, U. Tamponi, K. Tanida, F. Tenchini, F. Testa, A. Thaller, O. Tittel, R. Tiwary, E. Torassa, K. Trabelsi, F. F. Trantou, I. Tsaklidis, I. Ueda, T. Uglov, K. Unger, Y. Unno, K. Uno, S. Uno, P. Urquijo, Y. Ushiroda, S. E. Vahsen, R. van Tonder, K. E. Varvell, M. Veronesi, A. Vinokurova, V. S. Vismaya, L. Vitale, V. Vobbilisetti, R. Volpe, A. Vossen, E. Waheed, M. Wakai, S. Wallner, M.-Z. Wang, A. Warburton, S. Watanuki, C. Wessel, E. Won, X. P. Xu, B. D. Yabsley, S. Yamada, W. Yan, W. C. Yan, S. B. Yang, J. Yelton, K. Yi, J. H. Yin, K. Yoshihara, C. Z. Yuan, J. Yuan, L. Zani, F. Zeng, M. Zeyrek, B. Zhang, V. Zhilich, J. S. Zhou, Q. D. Zhou, L. Zhu, R. Žlebčík

https://belle.kek.jp/
https://www.belle2.org/

JHEP12(2025)109